ASYMPTOTIC EVALUATION OF THE PROBABILITIES OF MISCLASSIFICATION BY LINEAR DISCRIMINANT FUNCTIONS

BY

T. W. ANDERSON

TECHNICAL REPORT NO. 10
SEPTEMBER 28, 1972

PREPARED UNDER CONTRACT
N00014-67-A-0112-0030 (NR-042-034)
FOR THE OFFICE OF NAVAL RESEARCH

THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

ASYMPTOTIC EVALUATION OF THE PROBABILITIES OF MISCLASSIFICATION BY LINEAR DISCRIMINANT FUNCTIONS

BY

T. W. ANDERSON
Stanford University

TECHNICAL REPORT NO. 10

SEPTEMBER 28, 1972

PREPARED UNDER THE AUSPICES

OF

OFFICE OF NAVAL RESEARCH CONTRACT #N00014-67-A-0112-0030

THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

1. Introduction

The problem of classifying an observation into one of two multivariate normal populations with a common covariance matrix might be called the classical classification problem. Fisher's linear discriminant function [Fisher (1936)] serves as a criterion when samples are used to estimate the parameters of the two distributions. The exact probabilities of misclassifications when using this criterion are difficult to compute because the distribution of the criterion is virtually intractable. Wald (1944) made considerable progress towards finding the distribution, but only managed to express the criterion as a function of three angles whose distribution he gave. T. W. Anderson (1951) and Rosedith Sitgreaves (1952) continued with the problem. For further references see T. W. Anderson, Das Gupta, and Styan (1972), Subject Matter Code 6.2.

If the parameters are known, the Neyman-Pearson Fundamental Lemma can be applied to the classical classification problem [as done by Wald (1944)] to obtain a discriminant function that is linear in the components of the vector to be classified. The distribution of this statistic is normal; the mean and variance depends only on the Mahalanobis distance between the two populations. Since the procedure for classification is to classify into one population or the other depending on whether this statistic is greater or less than a constant, the probabilities of misclassification are found directly from the normal distribution. If the constant is 0, the probabilities are equal and the procedure is minimax.

*This paper was presented to the NATO Advanced Study Institute on Discriminant Analysis and Applications on June 12, 1972, at Kifissia, Greece.

When the parameters are unknown and there is available a sample from each population, the mean of each population is estimated by the mean of the respective sample and the common covariance matrix of the populations is estimated by using deviations from the respective means in the two samples. The classification function W, proposed by T. W. Anderson (1951), is obtained by replacing the parameters in the linear function resulting from the Neyman-Pearson Fundamental Lemma by the estimates; the substitution for parameters has been called "plugging in" estimates. This criterion differs from Fisher's discriminant function by subtraction of the average of the Fisher discriminant function at the two sample means. Then the distribution depends only on the population distance, and this fact makes the distribution problem simpler [T. W. Anderson (1951) and Sitgreaves (1952)], though it is still rather intractable.

When the sizes of the two samples increase, the limiting distribution of W approaches a normal distribution, whose mean and variance depend on the Mahalanobis distance; if the limiting mean is subtracted from W and the difference is divided by the limiting standard deviation, the statistic has the standard normal distribution as its limiting distribution. Bowker and Sitgreaves (1961) and Okamoto (1963) with correction (1968) have given asymptotic expansions of the distributions to the order of the reciprocal of the square of the sample sizes. The approximate probability depends on the unknown parameter (the distance).

The "Studentized" W statistic is W less the estimate of its limiting mean divided by the estimate of its limiting standard deviation. It, too, has the standard normal distribution as its limiting distribution. If a statistician wants to set his cut-off point to

achieve a specified probability of misclassification, he can use this Studentized W. An asymptotic expansion of the distribution of this statistic has been given by T. W. Anderson (1972).

In this paper we compare these two approximations to the probabilities of misclassification and their uses. For further discussion of the classification problem see Anderson (1958), Chapter 6.

2. The asymptotic expansion of the distribution of the classification statistic W

Let the two populations be $N(\mathbf{\mu}^{(1)}, \mathbf{\Sigma})$ and $N(\mathbf{\mu}^{(2)}, \mathbf{\Sigma})$, and let the two samples be $\mathbf{x}_1^{(1)}, \ldots, \mathbf{x}_{N_1}^{(1)}$ and $\mathbf{x}_1^{(2)}, \ldots, \mathbf{x}_{N_2}^{(2)}$, respectively. The observation to be classified is $\mathbf{x}$, which has the distribution $N(\mathbf{\mu}, \mathbf{\Sigma})$, where $\mathbf{\mu} = \mathbf{\mu}^{(1)}$ or $\mathbf{\mu} = \mathbf{\mu}_1^{(2)}$. The classification statistic W is

$$(1) \qquad W = (\bar{\mathbf{x}}^{(1)} - \bar{\mathbf{x}}^{(2)})' \mathbf{S}^{-1} [\mathbf{x} - \tfrac{1}{2} (\bar{\mathbf{x}}^{(1)} + \bar{\mathbf{x}}^{(2)})] ,$$

where

$$(2) \qquad \bar{\mathbf{x}}^{(1)} = \frac{1}{N_1} \sum_{j=1}^{N_1} \mathbf{x}_j^{(1)} , \quad \bar{\mathbf{x}}^{(2)} = \frac{1}{N_2} \sum_{j=1}^{N_2} \mathbf{x}_j^{(2)} ,$$

$$(3) \qquad n\mathbf{S} = \sum_{j=1}^{N_1} (\mathbf{x}_j^{(1)}-\bar{\mathbf{x}}^{(1)})(\mathbf{x}_j^{(1)}-\bar{\mathbf{x}}^{(1)})' + \sum_{j=1}^{N_2} (\mathbf{x}_j^{(2)}-\bar{\mathbf{x}}^{(2)})(\mathbf{x}_j^{(2)}-\bar{\mathbf{x}}^{(2)})' ,$$

and $n = N_1+N_2 - 2$. The rule is to classify $\mathbf{x}$ as coming from $N(\mathbf{\mu}^{(1)}, \mathbf{\Sigma})$ if $W > c$ and from $N(\mathbf{\mu}^{(2)}, \mathbf{\Sigma})$ if $W \leq c$, where c may be a constant, particularly 0, or a function of $\bar{\mathbf{x}}^{(1)}$, $\bar{\mathbf{x}}^{(2)}$, and $\mathbf{S}$.

The squared Mahalanobis distance is

$$(4) \qquad \alpha = (\underset{\sim}{\mu}^{(1)} - \underset{\sim}{\mu}^{(2)})' \underset{\sim}{\Sigma}^{-1} (\underset{\sim}{\mu}^{(1)} - \underset{\sim}{\mu}^{(2)}) ,$$

which can be estimated by

$$(5) \qquad a = (\bar{\underset{\sim}{x}}^{(1)} - \bar{\underset{\sim}{x}}^{(2)})' \underset{\sim}{S}^{-1} (\bar{\underset{\sim}{x}}^{(1)} - \bar{\underset{\sim}{x}}^{(2)}) .$$

The limiting distribution of W as $N_1 \to \infty$ and $N_2 \to \infty$ is normal with variance α and mean $\frac{1}{2}\alpha$ if $\underset{\sim}{x}$ is from $N(\underset{\sim}{\mu}^{(1)}, \underset{\sim}{\Sigma})$ and mean $-\frac{1}{2}\alpha$ if $\underset{\sim}{x}$ is from $N(\underset{\sim}{\mu}^{(2)}, \underset{\sim}{\Sigma})$; that is, the standard normal distribution $N(0, 1)$ is the limiting distribution of $(W - \frac{1}{2}\alpha)/\sqrt{\alpha}$ for $\underset{\sim}{x}$ coming from $N(\underset{\sim}{\mu}^{(1)}, \underset{\sim}{\Sigma})$ and of $(W + \frac{1}{2}\alpha)/\sqrt{\alpha}$ for $\underset{\sim}{x}$ coming from $N(\underset{\sim}{\mu}^{(2)}, \underset{\sim}{\Sigma})$.

Okamoto's expansion of the probability distribution [(1963), Corollary 1] to terms of order n^{-1} is

$$(6) \quad \Pr\left\{\frac{W - \frac{1}{2}\Delta^2}{\Delta} \le u \,\middle|\, \underset{\sim}{\mu} = \underset{\sim}{\mu}^{(1)}\right\} = \Phi(u) + \frac{1}{n}\phi(u)\left\{\Delta\left[1 + \frac{p}{2}k - \frac{p-2}{2}\frac{1}{k})\frac{1}{\Delta^2} - \frac{p-1}{2}\right]\right.$$

$$-\left[(1 + \frac{1}{2}k + \frac{1}{2}\frac{1}{k})\frac{p-3}{\Delta^2} + \frac{3p-2}{2} + \frac{1}{2}\frac{1}{k} + \frac{\Delta^2}{4}\right]u$$

$$\left.-\Delta\left[(1 + \frac{1}{k})\frac{1}{\Delta^2} + 1\right]u^2 - \left[(1 + \frac{1}{2}k + \frac{1}{2}\frac{1}{k})\frac{1}{\Delta^2} + 1\right]u^3\right\}$$

$$+ O(n^{-2}) ,$$

where $k = \lim_{n\to\infty} N_1/N_2$ as $N_1 \to \infty$ and $N_2 \to \infty$, $\Delta^2 = \alpha$, and $\Phi(\ )$ and $\phi(\ )$ are the cumulative distribution function and density of $N(0, 1)$, respectively. If $\lim_{n\to 1} N_1/N_2 = 1$, then

$$(7) \quad \Pr\left\{\frac{W - \frac{1}{2}\Delta^2}{\Delta} \le u \,\middle|\, \underset{\sim}{\mu} = \underset{\sim}{\mu}^{(1)}, \lim_{n\to\infty}\frac{N_1}{N_2} = 1\right\} = \Phi(u) + \frac{1}{n}\phi(u)\left\{\Delta\left[\frac{2}{\Delta^2} - \frac{p-1}{2}\right]\right.$$

$$\left.- 2\left[\frac{p-3}{\Delta^2} + \frac{3p-1}{2} + \frac{\Delta^2}{4}\right]u - \Delta\left[\frac{2}{\Delta^2} + 1\right]u^2 - \left[\frac{2}{\Delta^2} + 1\right]u^3\right\} + O(n^{-2})$$

$$= \Phi(u) + \frac{1}{n}\phi(u)\left\{\left[\frac{2}{\Delta^2}+1\right](\Delta+u)(1-u^2) - \frac{p+1}{2}\Delta\right.$$

$$\left. - \left[2\,\frac{p-2}{\Delta^2} + \frac{3p+1}{2} + \frac{\Delta^2}{4}\right]u\right\} + O(n^{-2}) \ .$$

The relation between the cut-off point c and the argument u is

(8) $$c = u\Delta + \frac{1}{2}\Delta^2 \ , \quad u = \frac{c - \frac{1}{2}\Delta^2}{\Delta} \ .$$

The probability of misclassification when $\underset{\sim}{x}$ is from $N(\underset{\sim}{\mu}^{(1)}, \underset{\sim}{\Sigma})$ is (6) [or (7)] with u given by (8); the probability depends importantly on the parameter .

A cut-off point of particular interest is $c = 0$, which corresponds to $u = -\frac{1}{2}\Delta$. If $N_1 = N_2$, this defines a minimax procedure. In this case the probability of misclassification is

(9) $$\Pr\left\{W \le 0 \,\middle|\, \underset{\sim}{\mu}=\underset{\sim}{\mu}^{(1)},\ \lim_{n\to\infty}\frac{N_1}{N_2} = 1\right\} = \Phi(-\frac{\Delta}{2}) + \frac{1}{n}\ \phi(\frac{\Delta}{2})\left\{\frac{p-1}{\Delta} + \frac{p}{4}\Delta\right\}$$
$$+ O(n^{-2}) \ .$$

As far as this approximation goes, the correction term is positive; that is, the probability of a misclassification error is greater than the value of the normal approximation. For a given value of Δ the correction term and hence the probability (to order n^{-1}) increases with p. For a given value of p the probability (to order n^{-1}) decreases with Δ.

Okamoto (as well as Bowker and Sitgreaves) expanded the characteristic function. The method of Anderson (1972) could be used to obtain the result.

3. The asymptotic expansion of the distribution of the Studentized W

To use the approximate probability given by (6) one must know the parameter $\alpha = \Delta^2$, but this is generally unknown; then the statistician cannot achieve, even approximately, a desired probability. However, he can use the fact that a is a consistant estimate of α and therefore $(W - \frac{1}{2} a)/\sqrt{a}$ and $(W + \frac{1}{2} a)/\sqrt{a}$ have $N(0, 1)$ as the limiting distribution in cases $\underset{\sim}{\mu} = \underset{\sim}{\mu}^{(1)}$ and $\underset{\sim}{\mu} = \underset{\sim}{\mu}^{(2)}$, respectively.

We can write

$$W - \frac{1}{2} a = (\bar{\underset{\sim}{x}}^{(1)} - \bar{\underset{\sim}{x}}^{(2)})' \underset{\sim}{S}^{-1} (\underset{\sim}{x} - \bar{\underset{\sim}{x}}^{(1)}) . \tag{10}$$

Then

$$\Pr\left\{\frac{W - \frac{1}{2} a}{\sqrt{a}} \leq u\right\} = \Pr\Big\{(\bar{\underset{\sim}{x}}^{(1)} - \bar{\underset{\sim}{x}}^{(2)})' \underset{\sim}{S}^{-1} (\underset{\sim}{x}-\underset{\sim}{\mu}) \tag{11}$$

$$\leq u \sqrt{(\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})'\underset{\sim}{S}^{-1}(\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})} + (\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})'\underset{\sim}{S}^{-1}(\bar{\underset{\sim}{x}}^{(1)}-\underset{\sim}{\mu})\Big\} .$$

Since $\underset{\sim}{x}$ has the distribution $N(\underset{\sim}{\mu}, \underset{\sim}{\Sigma})$ independently of $\bar{\underset{\sim}{x}}^{(1)}$, $\bar{\underset{\sim}{x}}^{(2)}$, and $\underset{\sim}{S}$, the conditional distribution of $(\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})'\underset{\sim}{S}^{-1}(\underset{\sim}{x}-\underset{\sim}{\mu})$ is $N[0, (\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})'\underset{\sim}{S}^{-1}\underset{\sim}{\Sigma}\underset{\sim}{S}^{-1}(\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})]$, and

$$r = \frac{(\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})'\underset{\sim}{S}^{-1}(\underset{\sim}{x}-\underset{\sim}{\mu})}{\sqrt{(\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})'\underset{\sim}{S}^{-1}\underset{\sim}{\Sigma}\underset{\sim}{S}^{-1}(\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})}} \tag{12}$$

has the distribution $N(0, 1)$. Then (11) is

$$(13)\qquad \Pr\left\{\frac{W-\frac{1}{2}a}{\sqrt{a}} \le u\right\} = \Pr\left\{r \le \frac{u\sqrt{(\bar{\mathbf{x}}^{(1)}-\bar{\mathbf{x}}^{(2)})'\mathbf{S}^{-1}(\bar{\mathbf{x}}^{(1)}-\bar{\mathbf{x}}^{(2)})}+(\bar{\mathbf{x}}^{(1)}-\bar{\mathbf{x}}^{(2)})'\mathbf{S}^{-1}(\bar{\mathbf{x}}^{(1)}-\boldsymbol{\mu})}{\sqrt{(\bar{\mathbf{x}}^{(1)}-\bar{\mathbf{x}}^{(2)})'\mathbf{S}^{-1}\boldsymbol{\Sigma}\mathbf{S}^{-1}(\bar{\mathbf{x}}^{(1)}-\bar{\mathbf{x}}^{(2)})}}\right\}$$

$$= \mathscr{E}\,\Phi\left[\frac{u\sqrt{(\bar{\mathbf{x}}^{(1)}-\bar{\mathbf{x}}^{(2)})'\mathbf{S}^{-1}(\bar{\mathbf{x}}^{(1)}-\bar{\mathbf{s}}^{(2)})}+(\bar{\mathbf{x}}^{(1)}-\bar{\mathbf{x}}^{(2)})'\mathbf{S}^{-1}(\bar{\mathbf{x}}^{(1)}-\boldsymbol{\mu})}{(\bar{\mathbf{x}}^{(1)}-\bar{\mathbf{x}}^{(2)})'\mathbf{S}^{-1}\boldsymbol{\Sigma}\mathbf{S}^{-1}(\bar{\mathbf{x}}^{(1)}-\bar{\mathbf{x}}^{(2)})}\right],$$

where the expectation is with respect to $\bar{\mathbf{x}}^{(1)}$, $\bar{\mathbf{x}}^{(2)}$, and $\mathbf{S}$.

When $\boldsymbol{\mu} = \boldsymbol{\mu}^{(1)}$, $\bar{\mathbf{x}}^{(1)} - \bar{\mathbf{x}}^{(2)}$, $\bar{\mathbf{x}}^{(1)} - \boldsymbol{\mu}$, and $\mathbf{S}$ converge in probability to $\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)}$, $\mathbf{0}$, and $\boldsymbol{\Sigma}$, respectively. We can expand the argument of $\Phi(\ )$ in a Taylor's series in terms of $\sqrt{n}$ times the differences between the estimates and their probability limits. When the expansion includes third degree terms and the expecations computed, the result is

$$(14)\qquad \Pr\left\{\frac{W-a}{\sqrt{a}} \le u \,\middle|\, \boldsymbol{\mu}=\boldsymbol{\mu}^{(1)}\right\} = \Phi(u) + \frac{1}{n}\,\phi(u)\left[\frac{(p-1)}{\Delta}(1+k) - (p-\frac{1}{4}+\frac{1}{2}k)u - \frac{1}{4}u^3\right] + O(n^{-2})$$

Interchanging N_1 and N_2 gives

$$(15)\qquad \Pr\left\{\frac{W+\frac{1}{2}a}{\sqrt{a}} \le v \,\middle|\, \boldsymbol{\mu}=\boldsymbol{\mu}^{(2)}\right\} = \Phi(v) - \frac{1}{n}\,\phi(v)\left[\frac{p-1}{\Delta}(1+\frac{1}{k}) + (p-\frac{1}{4}+\frac{1}{2k})v + \frac{1}{4}v^3\right] + O(n^{-2})\ .$$

The proof of these results was given by T. W. Anderson (1972). If $\lim_{n\to\infty} N_1/N_2 = k = 1$,

$$(16)\qquad \Pr\left\{\frac{W-a}{\sqrt{a}} \le u \,\middle|\, \boldsymbol{\mu}=\boldsymbol{\mu}^{(1)},\ \lim_{n\to\infty}\frac{N_1}{N_2} = 1\right\} = \Phi(u) + \frac{1}{n}\,\phi(u)\left\{2\,\frac{p-1}{\Delta} - (p+\frac{1}{4})u - \frac{1}{4}u^3\right\} + O(n^{-2})\ .$$

The correction term in (14) [(15) or (16)] is positive for $u < 0$. If $p = 1$, the correction term does not depend on Δ; if $p > 1$, the correction term decreases with Δ. For $u < 0$, the correction term increases with p.

For $u = -\frac{1}{2}\Delta$ (which is not $c = 0$)

$$(17) \qquad \Pr\left\{\frac{W-a}{\sqrt{a}} \leq -\frac{\Delta}{2} \Big| \underset{\sim}{\mu} = \underset{\sim}{\mu}^{(1)},\ \lim_{n\to\infty}\frac{N_1}{N_2} = 1\right\} = \Phi(-\frac{\Delta}{2}) + \frac{1}{n}\phi(\frac{\Delta}{2})\left\{2\,\frac{p-1}{\Delta} + \frac{4p+1}{8}\,\Delta + \frac{\Delta^3}{32}\right\} + O(n^{-2})\ .$$

4. Numerical values of the correction term for the Studentized W when $N_1 = N_2$

We can obtain an idea of the importance of the term of order $1/n$ by studying numerical values of it. We consider the second term in (16), which is the error to order n^{-1} of using $\Phi(u)$ for the probability of misclassification. The correction relative to the nominal probability of misclassification is

$$(18) \qquad \frac{1}{n}\ \frac{\phi(u)}{\Phi(u)}\left[2\,\frac{p-1}{\Delta} - (p + \frac{1}{4})\,u - \frac{u^3}{4}\right].$$

Table 1 gives values of the term in brackets for the five values of u corresponding to values of $\Phi(u)$ of .1, .05, .025, .01, and .005, and various values of p and Δ. It is 4.0893 for $u = -1.28155$ [$\Phi(u) = .1$], $p = 2$, and $\Delta = 2$. The correction relative to the nominal probability of misclassification is the value in the table multiplied by the ratio $\phi(u)/\Phi(u)$ divided by $n = N_1 + N_2 - 2$. In the example above it is $4.0893 \times 1.755 = 7.1767$ divided by n. If $N_1 = N_2 = 25$, then $n = 48$ and the correction relative to the nominal probability of misclassification

is about .15. Here the correction would be rather small. For values of N_1 and N_2 somewhat larger, one might be willing to neglect the correction. One would hope that for these values of N_1 and N_2 the error when using this correction term would be rather small.

We might also be interested in the correction at $u = -\frac{1}{2}\Delta$. Table 2 gives the information. For example, for $\Delta = 4$ $\Phi(-\frac{1}{2}\Delta) = .022750$ (which would be the minimax probability if the parameters were known) and the correction is the appropriate number in the fourth column multiplied by .053991 divided by n. If $N_1 = N_2 = 25$ and $p = 2$, then $n = 48$ and the correction relative to the nominal probability is $7 \times 2.383/48 = .3475$.

5. Comparison of the expansions of the distributions of W and the Studentized W

It is striking that the asymptotic expansion of the distribution of the Studentized W is much simpler than that of W itself [the comparison of (6) with (14) and (7) with (16)], except for the particular case of $u = -\frac{1}{2}\Delta$ [(9) with (17)] which has special meaning for W $(c = 0)$, but not for the Studentized W.

It is of interest to compare the correction terms of the two asymptotic expansions. The difference is

$$(19)\quad \Pr\left\{\frac{W - \frac{1}{2}a}{\sqrt{a}} \le u \middle| \underset{\sim}{\mu} = \underset{\sim}{\mu}^{(1)}\right\} - \Pr\left\{\frac{W - \frac{1}{2}\alpha}{\sqrt{\alpha}} \le u \middle| \underset{\sim}{\mu} = \underset{\sim}{\mu}^{(1)}\right\} = \frac{1}{n}\phi(u)\left\{\frac{p-2}{2}\,\frac{2+k+1/k}{\Delta} + \frac{p-1}{2}\Delta \right.$$

$$+ \left[(1 + \frac{1}{2}k + \frac{1}{2}\frac{1}{k})\frac{p-3}{\Delta^2} + \frac{2p-3}{4} - \frac{1}{2}k + \frac{1}{2}\frac{1}{k} + \frac{\Delta^2}{4}\right] u$$

$$\left. + \left[(1 + \frac{1}{k})\frac{1}{\Delta} + \Delta\right] u^2 + \left[\frac{2+k+1/k}{2\Delta^2} + \frac{3}{4}\right] u^3 \right\} + O(n^{-2}) .$$

If $\lim_{n\to\infty} N_1/N_2 = k = 1$, the expression simplifies to

$$(20) \quad \Pr\left\{\frac{W - \frac{1}{2}a}{\sqrt{a}} \le u \,\middle|\, \underset{\sim}{\mu} = \underset{\sim}{\mu}^{(1)},\ \lim_{n\to\infty} \frac{N_1}{N_2} = 1\right\} - \Pr\left\{\frac{W - \frac{1}{2}\alpha}{\sqrt{\alpha}} \le u \,\middle|\, \underset{\sim}{\mu} = \underset{\sim}{\mu}_1,\ \lim_{n\to\infty} \frac{N_1}{N_2} = 1\right\}$$

$$= \frac{1}{n}\phi(v)\left\{2\,\frac{p-2}{\Delta} + \frac{p-1}{2}\Delta + 2\left[\frac{p-3}{\Delta^2} + \frac{2p-3}{4} + \frac{\Delta^2}{4}\right]u + \left[\frac{2}{\Delta} + \Delta\right]u^2 - \left[\frac{2}{\Delta^2} + \frac{3}{4}\right]u^3\right\} + O(n^{-2}) .$$

In particular, for $u = -\frac{1}{2}\Delta$ the difference is

$$(21) \quad \Pr\left\{\frac{W - \frac{1}{2}a}{\sqrt{a}} \le -\frac{\Delta}{2} \,\middle|\, \underset{\sim}{\mu} = \underset{\sim}{\mu}^{(1)},\ \lim_{n\to\infty} \frac{N_1}{N_2}\right\} - \Pr\left\{W \le 0 \,\middle|\, \underset{\sim}{\mu} = \underset{\sim}{\mu}^{(1)},\ \lim_{n\to\infty} \frac{N_1}{N_2} = 1\right\}$$

$$= \frac{1}{n}\phi\left(\frac{\Delta}{2}\right)\left\{\frac{p-1}{\Delta} + \left(\frac{p}{4} + \frac{1}{8}\right)\Delta + \frac{1}{32}\Delta^3\right\} + O(n^{-2}) .$$

Put another way, the correction term for $\Pr\{(W-a)/\sqrt{a} \le -\frac{1}{2}\Delta\}$ is twice the correction term for $\Pr\{W \le 0\}$ plus $\phi(\frac{1}{2}\Delta)\{\Delta/8 + \Delta^3/32\}/n$. The latter term, which does not depend on p, is usually small; values of $\Delta/8 + \Delta^3/32$ are given in Table 3. Comparison with Table 2 shows that for $p > 1$ this term is small except for large Δ. Thus, roughly speaking, the correction for the Studentized W is about that of W itself.

Okamoto (1963) has given numerical values of the term of order $1/n$ and the term of order $1/n^2$ in the expansion of $\Pr\{W \le 0 | \underset{\sim}{\mu} = \underset{\sim}{\mu}^{(1)}\}$ for $N_1 = N_2 = 100$ $(n = 198)$ for various values of p and Δ. His values for $1/n$ are about twice the values we can compute from Table 2. In his table for small values of p and Δ the ratio of the term of order $1/n^2$ to the term of order $1/n$ is very roughly $1/n$. The maximum of the $1/n^2$ term over Δ increases with p. At $p = 7$, for example, it is about .0008. The table suggests that for small or moderate values of p the second correction term can be safely ignored for moderately large values of N_1 and N_2.

6. Comparison of approximate densities and moments

Corresponding to the approximate distributions of $(W-\alpha)/\sqrt{\alpha}$ and $(W-a)/\sqrt{a}$ (for $\underset{\sim}{\mu}=\underset{\sim}{\mu}^{(1)}$) are densities and moments. It is of some interest to compare these.

The approximate density of $(W - \frac{1}{2}\Delta^2)/\Delta$ is

$$(22) \qquad \phi(u) \Bigg\{1 - \frac{1}{n}\Bigg[(1 + \frac{1}{2}k + \frac{1}{2}\frac{1}{k})\frac{p-3}{\Delta^2} + \frac{3p-2}{2} + \frac{1}{2}\frac{1}{k} + \frac{\Delta^2}{4}$$

$$+ \left(\frac{3 + \frac{1}{2}p\,k - \frac{1}{2}(p-6)/k}{\Delta} - \frac{p-5}{2}\Delta\right) u$$

$$- \left(\frac{p-6}{\Delta^2}(1 + \frac{1}{2}k + \frac{1}{2}\frac{1}{k}) + \frac{3p-8}{2} + \frac{1}{2}\frac{1}{k} + \frac{\Delta^2}{4}\right) u^2$$

$$+ \left(\frac{1 + 1/k}{\Delta} + \Delta\right) u^3 + \left(\frac{1 + \frac{1}{2}k + \frac{1}{2}\frac{1}{k}}{\Delta^2} + 1\right) u^4\Bigg]\Bigg\} ,$$

which for $k = 1$ is

$$(23) \qquad \phi(u) \Bigg\{1 - \frac{1}{n}\Bigg[2\frac{p-3}{\Delta^2} + \frac{3p-1}{2} + \frac{\Delta^2}{4} + \left(\frac{6}{\Delta} - \frac{p-5}{2}\Delta\right) u - \left(2\frac{p-6}{\Delta^2} + \frac{3p-7}{2} + \frac{\Delta^2}{4}\right)u^2$$

$$+ \left(\frac{2}{\Delta} + \Delta\right) u^3 + \left(\frac{2}{\Delta^2} + 1\right) u^4\Bigg]\Bigg\} .$$

The approximate density of $(W - \frac{1}{2}a)/\sqrt{a}$ is

$$(24) \quad \phi(u) \left\{1 - \frac{1}{n}\left[p - \frac{1}{4} + \frac{1}{2}k + \frac{(p-1)(1+k)}{\Delta} u - (p - 1 + \frac{1}{2}k)u^2 - \frac{u^4}{4}\right]\right\} ,$$

which for $k = 1$ is

$$(25) \qquad \phi(u) \left\{1 - \frac{1}{n}\left[p + \frac{1}{4} + 2\frac{p-1}{\Delta} u - (p - \frac{1}{2}) u^2 - \frac{u^4}{4}\right]\right\} .$$

The approximate mean of $(W - \frac{1}{2}\Delta^2)/\Delta$ is

$$(26) \qquad -\frac{1}{n}\left[\frac{6 + \frac{1}{2}pk - \frac{1}{2}(p-12)/k}{\Delta} - \frac{p-11}{2}\Delta\right],$$

which for $k = 1$ is

$$(27) \qquad -\frac{1}{n}\left[\frac{12}{\Delta} - \frac{p-11}{2}\Delta\right];$$

the approximate second-order moment is

$$(28) \qquad 1 + \frac{1}{n}\left[\frac{(2p-30)\ (1 + \frac{1}{2}k + \frac{1}{2}\frac{1}{k})}{\Delta^2} + 3p + 26 - \frac{1}{k} + \frac{1}{2}\Delta^2\right],$$

which for $k = 1$ is

$$(29) \qquad 1 + \frac{1}{n}\left[\frac{4p - 60}{\Delta^2} + 3p - 25 + \frac{1}{2}\Delta^2\right].$$

The approximate mean of $(W - \frac{1}{2}a)/\sqrt{a}$ is

$$(30) \qquad -\frac{1}{n}\frac{(p-1)(1+k)}{\Delta},$$

which for $k = 1$ is

$$(31) \qquad -\frac{1}{n}\,2\,\frac{p-1}{\Delta};$$

the approximate second-order moment is

$$(32) \qquad 1 + \frac{1}{n}(2p + 1 + k),$$

which for $k = 1$ is

$$(33) \qquad 1 + \frac{1}{n}(2p + 2).$$

In each case the "approximate" moment is the moment of the approximate density. The approximate second-order moment is also the approximate variance. For $(W - \frac{1}{2} a)/\sqrt{a}$ the approximate mean is negative for $p > 1$ (while it is 0 for the standard normal distribution); its numerical value increases with p and decreases with Δ. The approximate variances are greater than 1 (the value for the standard normal distribution); it increases with p, but does not depend on Δ.

7. Achieving a given probability of misclassification

Suppose one wants to achieve a given probability p of misclassification when $\underset{\sim}{\mu}=\underset{\sim}{\mu}^{(1)}$, say. How should one choose the cut-off point $c = u\sqrt{a} + \frac{1}{2} a$ for W or equivalently u for $(W - \frac{1}{2} a)/\sqrt{a}$?

Let u_0 be the number such that $\Phi(u_0) = p$. Then the probability of misclassification is

$$(34) \quad p + \frac{1}{n} \phi(u_0) \left[\frac{(p-1)(1+k)}{\Delta} - (p - \frac{1}{4} + \frac{1}{2} k)u_0 - \frac{1}{4} u_0^3\right] + O(n^{-2}) \quad .$$

The correction term of order n^{-1} contains the unknown parameter Δ (if $p > 1$). However, Δ can be estimated by $\sqrt{a}$. These facts suggest taking

$$(35) \qquad u = u_0 - \frac{1}{n} \left[\frac{(p-1)(1+k)}{\sqrt{a}} - (p - \frac{1}{4} + \frac{1}{2} k) u_0 - \frac{1}{4} u_0^3\right] .$$

Then

$$(36) \quad \Pr\left\{\frac{W - \frac{1}{2} a}{\sqrt{a}} \leq u \,\middle|\, \underset{\sim}{\mu}=\underset{\sim}{\mu}^{(1)}\right\} = \Pr\left\{\frac{W - \frac{1}{2} a}{\sqrt{a}} + \frac{1}{n} \frac{(p-1)(1+k)}{\sqrt{a}} \leq u^*\right\} ,$$

where

$$(37) \qquad u^* = u_0 + \frac{1}{n}\left[\left(p - \frac{1}{4} + \frac{1}{2}k\right) u_0 + \frac{1}{4} u_0^3\right] .$$

If $p = 1$, this probability is (14) with $u = u^*$, which is $p+O(n^{-2})$. When $p > 1$, we calculate the probability of misclassification as

$$(38) \qquad \Pr\{W - \frac{1}{2} a \leq u^* \sqrt{a} - \frac{1}{n}(p-1)(1+k)\} = \Pr\{(\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})'\underset{\sim}{S}^{-1}(\underset{\sim}{x}-\underset{\sim}{\mu})$$

$$\leq u^* \sqrt{(\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})'\underset{\sim}{S}^{-1}(\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})} + (\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})'\underset{\sim}{S}^{-1}(\bar{\underset{\sim}{x}}^{(1)}-\underset{\sim}{\mu}) - \frac{1}{n}(p-1)(1+k)\}$$

$$= \Phi\left[\frac{u^* \sqrt{(\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})'\underset{\sim}{S}^{-1}(\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})} + (\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})'\underset{\sim}{S}^{-1}(\bar{\underset{\sim}{x}}^{(1)}-\underset{\sim}{\mu}) - \frac{1}{n}(p-1)(1+k)}{\sqrt{(\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})'\underset{\sim}{S}^{-2}(\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)})}}\right],$$

where $\bar{\underset{\sim}{x}}^{(1)}-\bar{\underset{\sim}{x}}^{(2)}$, $\bar{\underset{\sim}{x}}^{(1)}-\underset{\sim}{\mu}$ and $\underset{\sim}{S}$ have the joint distribution given in Anderson (1972). Then the expansion of $\Phi(\ )$ is

$$(39) \qquad \Phi\{u^* + \frac{1}{\sqrt{n}} C^*(\underset{\sim}{Z},\underset{\sim}{V}) + \frac{1}{n} D^*(\underset{\sim}{Y},\underset{\sim}{Z},\underset{\sim}{V}) + r^*_{7n}(\underset{\sim}{Y},\underset{\sim}{Z},\underset{\sim}{V})$$

$$- \frac{1}{n}(p-1)(1+k)\left[\frac{1}{\Delta} - \frac{1}{\Delta^3\sqrt{n}}(\underset{\sim}{\delta}'v_1-\underset{\sim}{\delta}\underset{\sim}{V}\underset{\sim}{\delta}) + r^*(\underset{\sim}{Y},\underset{\sim}{Z},\underset{\sim}{V})\right]$$

$$= \Phi(u^*) + \phi(u^*)\left\{\frac{1}{\sqrt{n}} C^*(\underset{\sim}{Z},\underset{\sim}{V}) + \frac{1}{n}\left[D^*(\underset{\sim}{Y},\underset{\sim}{Z},\underset{\sim}{V})\right.\right.$$

$$\left. - \frac{1}{2} u^* C^{*2}(\underset{\sim}{Z},\underset{\sim}{V}) - \frac{1}{n}(p-1)(1+k)\frac{1}{\Delta}\right]$$

$$\left. + \frac{1}{\Delta^3 n^{3/2}}(p-1)(1+k)(\underset{\sim}{\delta}'\underset{\sim}{Y}-\underset{\sim}{\delta}'\underset{\sim}{V}\underset{\sim}{\delta})\right\} + \frac{1}{n^{3/2}} r^*_8(\underset{\sim}{Y},\underset{\sim}{Z},\underset{\sim}{V}) + \frac{1}{n^2} r^*_9(\underset{\sim}{Y},\underset{\sim}{Z},\underset{\sim}{V})$$

$$+ r^*_{10n}(\underset{\sim}{Y},\underset{\sim}{Z},\underset{\sim}{V}) ,$$

where $C^*(\underset{\sim}{Z},\underset{\sim}{V})$, $D^*(\underset{\sim}{Y},\underset{\sim}{Z},\underset{\sim}{V})$, and $r^*_{7n}(\underset{\sim}{Y},\underset{\sim}{Z},\underset{\sim}{V})$ are $C(\underset{\sim}{Z},\underset{\sim}{V})$, $D(\underset{\sim}{Y},\underset{\sim}{Z},\underset{\sim}{V})$ and $r_{7n}(\underset{\sim}{Y},\underset{\sim}{Z},\underset{\sim}{V})$ of Anderson (1972), with u replaced by u^* and $r^*(\underset{\sim}{Y},\underset{\sim}{Z},\underset{\sim}{V})$ in the remainder term in (19) of Anderson (1972). The expected value of $\Phi(\ )$ is

$$(40) \qquad \Phi(u^*) + \frac{1}{n}\,\phi(u^*)\left[-\left(p - \frac{1}{4} + \frac{1}{2k}\right) u^* - \frac{1}{4}\, u^{*3}\right] + O(n^{-2})$$

$$= \Phi(u_0) + O(n^{-2})$$

$$= p + O(n^{-2}) \ .$$

TABLE 1

$$2\,\frac{p-1}{\Delta} - \left(p + \frac{1}{4}\right)u - \frac{u^3}{4}$$

	p \ Δ	1	2	3	4	6	∞
$u = -1.28155$	1	2.13	2.13	2.13	2.13	2.13	2.13
$\phi(u) = .100$	2	5.41	4.41	4.08	3.91	3.74	3.41
$\Phi(u) = .17550$	4	11.97	8.97	7.97	7.47	6.97	5.97
$\phi(u)/\Phi(u) = 1.755$	8	25.10	18.10	15.77	14.60	13.43	11.10

	p \ Δ	1	2	3	4	6	∞
$u = -1.64485$	1	3.17	3.17	3.17	3.17	3.17	3.17
$\Phi(u) = .05$	2	6.81	5.81	5.48	5.31	5.15	4.81
$\phi(u) = .10314$	4	14.10	11.10	10.10	9.50	9.10	8.10
$\phi(u)/\Phi(u) = 2.063$	8	28.68	21.68		18.18	17.06	14.68

	p \ Δ	1	2	3	4	6	∞
$u = -1.95996$	1	4.33	4.33	4.33	4.33	4.33	4.33
$\Phi(u) = .025$	2	8.29	7.29	6.96	6.79	6.63	6.29
$\phi(u) = .05844$	4	16.21	13.21	12.21	11.71	11.21	10.21
$\phi(u)/\Phi(u) = 2.338$	8	32.05	25.05	22.72	21.55	20.39	18.05

	p \ Δ	1	2	3	4	6	∞
$u = -2.32635$	1	6.06	6.06	6.06	6.06	6.06	6.06
$\phi(u) = .01$	2	10.38	9.38	9.05	8.88	8.72	8.38
$\Phi(u) = .02665$	4	19.03	16.03	15.03	14.53	14.03	13.03
$\phi(u)/\Phi(u) = 2.665$	8	36.34	29.34	27.01	25.84	24.67	22.34

	p \ Δ	1	2	3	4	6	∞
$u = -2.57583$	1	7.49	7.49	7.49	7.49	7.49	7.49
$\phi(u) = .005$	2	11.07	11.07	10.73	10.57	10.40	10.07
$\Phi(u) = .01446$	4	18.22	18.22	17.22	16.72	16.22	15.22
$\phi(u)/\Phi(u) = 2.892$	8	32.52	32.52	30.19	29.02	27.86	25.52

TABLE 2

$$2\,\frac{p-1}{\Delta} + \left(\frac{p}{2} + \frac{1}{8}\right) + \frac{\Delta^3}{32}$$

p \ Δ	1	2	3	4	6
1	.65625	1.50000	2.71875	4.50000	10.50000
2	3.15625	3.50000	4.88542	7.00000	13.83333
4	8.15125	7.50000	9.21875	12.00000	20.00000
8	18.15625	15.50000	17.88542	22.00000	40.83333
$\phi(-\frac{1}{2}\Delta)$	.35206	.24197	.129518	.053991	.0044318
$\Phi(-\frac{1}{2}\Delta)$	.30854	.15866	.066807	.022750	.0013499
$\phi(-\frac{1}{2}\Delta)/\Phi(\frac{1}{2}\Delta)$	1.141	1.525	1.939	2.383	3.283

TABLE 3

$$\frac{\Delta}{8} + \frac{\Delta^3}{32}$$

Δ	1	2	3	4	6
$\frac{\Delta}{8} + \frac{\Delta^2}{32}$	.15626	.50000	1.21875	2.50000	7.50000

REFERENCES

[1] Anderson, T. W. (1951), Classification by multivariate analysis, Psychometrika, 16, 31-50.

[2] Anderson, T. W. (1958), An Introduction to Multivariate Statistical Analysis, John Wiley & Sons, Inc., New York.

[3] Anderson, T. W. (1972), An asymptotic expansion of the distribution of the "Studentized" classification statistic W, Technical Report No. 9, Stanford University.

[4] Anderson, T. W., Somesh Das Gupta and George P. H. Styan (1972), A Bibliography of Multivariate Statistical Analysis, Oliver & Boyd, Ltd., Edinburgh.

[5] Bowker, Albert H. and Rosedith Sitgreaves (1961), An asymptotic expansion for the distribution function of the W-classification statistic, Stud. Item Anal. Predict. (H. Solomon, ed.) 285-292.

[6] Fisher, Ronald A. (1936), The use of multiple measurements in taxonomic problems, Ann. Eugenics, 7, 179-188.

[7] Okamoto, Mashashi (1963), An asymptotic expansion for the distribution of linear discriminant function, Ann. Math. Statist., 34, 1286-1301.

[8] Sitgreaves, Rosedith (1952), On the distribution of two random matrices used in classification procedures, Ann. Math. Statist., 23, 263-270.

[9] Wald, Abraham (1944), On a statistical problem arising in the classification of an individual into one of two groups, Ann. Math. Statist., 15, 145-162.

OTHER REPORTS IN THIS SERIES

1. Confidence Limits for the Expected Value of an Arbitrary Bounded Random Variable with a Continuous Distribution Function by T. W. Anderson.

2. Efficient Estimation of Regression Coefficients in Time Series by T. W. Anderson.

3. Determining the Appropriate Sample Size for Confidence Limits for a Proportion by T. W. Anderson and Herman Burstein.

4. Some General Results on Time-Ordered Classification by D. V. Hinkley.

5. Tests for Randomness of Directions Against Equatorial and Bimodal Alternatives by T. W. Anderson and M. A. Stephens.

6. Estimation of Covariance Matrices with Linear Structure and Moving Average Processes of Finite Order by T. W. Anderson.

7. The Stationarity of an Estimated Autoregressive Process by T. W. Anderson.

8. On the Inverse of Some Covariance Matrices of Toeplitz Type by Raul Pedro Mentz.

9. An Asymptotic Expansion of the Distribution of the "Studentized" Classification Statistic W by T. W. Anderson.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Department of Statistics Stanford University Stanford, California 94305	2a. REPORT SECURITY CLASSIFICATION 2b. GROUP

3. REPORT TITLE
Asymptotic Evaluation of the Probabilities of Misclassification by Linear Discriminant Functions

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Technical Report

5. AUTHOR(S) (Last name, first name, initial)
T. W. Anderson

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
September 28, 1972	19	9

8a. CONTRACT OR GRANT NO. N-000014-67-A-0112-0030 b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 10 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. AVAILABILITY/LIMITATION NOTICES

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research

13. ABSTRACT.

Linear discriminant functions are used to classify an observation as coming from one of two normal populations with common covariance matrices and different means when samples are used to estimate the parameters of the distributions. Okamoto's asymptotic expansion of the distribution of the classification statistic W is compared with Anderson's expansion for the Studentized W (that is, W standardized by estimates of its mean and standard deviation). Some numerical evaluations of the term of order of the reciprocal of the sample sizes is given. The uses of the two approximate distributions are discussed.

DD FORM 1 JAN 64 1473

UNCLASSIFIED
Security Classification

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
classification classification statistic asymptotic expansion discriminant analysis						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through
____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through
____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through
____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS), (S), (C), or (U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.